INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(54) Title: MULTICAST–ENABLED ADDRESS RESOLUTION PROTOCOL (ME–ARP)
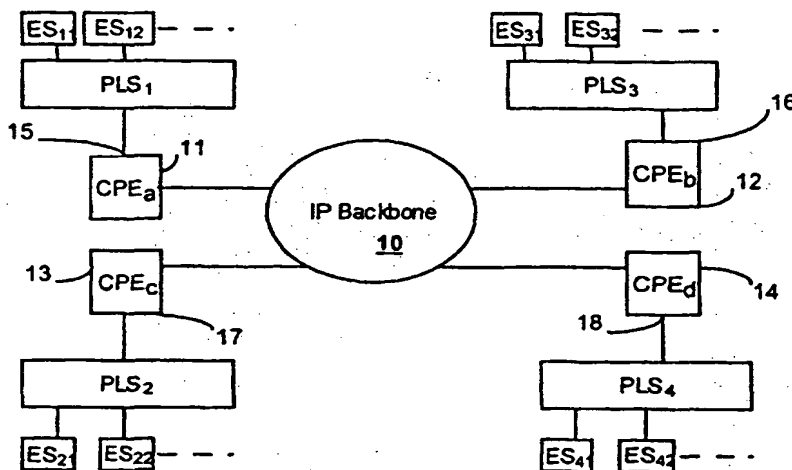
(57) Abstract

A Multicast–Enabled Address Resolution Protocol (ME–ARP) is disclosed. This ME–ARP allows the building of independent IP based Virtual Private LAN segments (VPLS) over a multicast enabled IP backbone using stateless tunnels and optimal VPLS traffic forwarding. Each VPLS has an associated IP subnet which is completely independent from other VPLS or the underlying IP backbone itself. Each Customer Premises Equipment (CPE) device needs only to be configured with a VPLS identifier and its serving IP subnet per VPLS designated interface.

MULTICAST-ENABLED ADDRESS
RESOLUTION PROTOCOL (ME-ARP)

FIELD OF THE INVENTION

5

This invention relates to a scalable and server-less solution to build Virtual Private LAN Segments (VPLS) based on a multicast enabled IP backbone and more particularly to a Multicast-Enabled Address Resolution Protocol (ME-ARP).

10   BACKGROUND OF THE INVENTION

The popularity of the Internet is driving requirements for secure and segregated IP interconnection of remote sites. One solution is to use the underlying network supporting virtual connections i.e. Frame Relay or ATM. These virtual connections can be separated

15   by provisioning to form a Virtual Private Network which is Layer 3 protocol transparent. However if the underlying network is IP itself, as is the case with the Internet then IP tunnels can be used to interconnect two or more sites. Any other known layer 2 VPN (Virtual Private Network) solution used in the prior art requires a centralized server where all CPE (Customer Premises Equipment) and IP devices have to be statically or

20   dynamically registered, like LANE (Local-Area-Network Emulation), NHRP (Next-Hop-Routing-Protocol) or Classical IP.

A need exists for building IP based virtual private LAN segments (sharing one IP subnet) with complete transparency regarding TCP/IP, site-independent CPE

25   configuration and with dynamic stateless tunnels to optimally forward unicast traffic based on routing and policy per VPLS. VPLS with different Identifiers can use overlapping IP subnets. With the method of the present invention, a centralized server or a list of CPE devices configured for each VPN is not required.

30   SUMMARY OF THE INVENTION

1

One aspect of the present invention is to provide a scalable and server-less solution to build Virtual Private LAN Segments (VPLS).

Another aspect of the present invention is to provide a Multicast-Enabled Address Resolution Protocol (ME-ARP). This invention allows the building of independent IP based Virtual Private LAN segments (VPLS) over a multicast enabled IP backbone using stateless tunnels and optimal VPLS traffic forwarding. Each VPLS has an associated IP subnet which is independent from other VPLS or the underlying IP backbone itself. Each Customer Premises Equipment (CPE) device needs only to be configured with a VPLS identifier and its serving IP subnet per VPLS designated interface. In addition, each end station connected to a Physical LAN Segment (PLS) does not need to be modified in order to be a member of the VPLS. No other configuration parameters e.g. list of CPE devices, their logical or physical locations nor their IP addresses are required. The unique invention is ME-ARP (Multicast Enabled Address Resolution Protocol) including the creation of constructed lower layer address based on VPN (Virtual Private Network) Id and tunnel endpoint. Advantages provided by the method of the present invention include:

a)   separation of customer IP address space from either the service provider or another customer determined by policy not to be in the same virtual private network (VPN);

b)   capability for a remote site to belong to one or more VPN as long as the VPN policy allows. To provide support for IPv4 based applications at this point;

c)   transparent or Routed VPN's (by use of external routers) can be constructed independently or combined with this architecture;

d)   due to the use of an underlying IP multicast network to forward VPN broadcast traffic in this solution ,there is no need to provide address or broadcast servers; and

e)   VPN traffic forwarding is achieved via stateless and optionally secured tunnels which are optimally routed using the underlying IP network backbone routing architecture.   2

BRIEF DESCRIPTION OF THE DRAWINGS

.FIG. 1a is a block diagram illustrating a physical view of a Virtual Private LAN Segment (VPLS) network for use with the present invention;

FIG. 1b is a diagram illustrating a logical view of the network of FIG. 1a or as would be seen from the customer's perspective;

FIG. 2a illustrates a packet format corresponding to an IPsec Authentication Header (AH) encapsulation with authentication;

FIG. 2b illustrates a packet format corresponding to an IPsec Encapsulating Security Payload (ESP) with authentication privacy;

FIG. 3 illustrates a standard ARP packet format on Ethernet;

FIG. 4 is a block diagram of a IP Backbone network for illustrating the ME-ARP request/reply packet flow according to the present invention;

Fig. 5 is a block diagram illustrating the transfer of ME-ARP packet information between a first and second end station according to the method of the present invention; and

Fig. 6 is a table illustrating the content of the ARP tables at various point during the transfer of ME-ARP packet information.

Similar references are used in different figures to denote similar components.

In order to facilitate the description of the invention, the following abbreviations have been used. The terminology used in this document is based on the definitions proposed by the Internet Engineers Task Force (IETF).

CBT   Core Based Tree Multicast Routing Protocol

SUBSTITUTE SHEET (RULE 26)

CPE	Customer Premises Equipment
DVMRP	Distance Vector Multicasting Routing Protocol
GRE	Generic Routing Encapsulation
IGMP	Internet Group Management Protocol
LAN	Local Area Network
MOSPF	Multicast extensions for Open Shortest Path First
PA	Provider Address
PIM	Protocol Independent Multicast
PLS	Physical LAN Segment
VPN	Virtual Private Network
VPLS	Virtual Private LAN
UVIP	Unnumbered VPN Internet Protocol

The term "Client Address" (CA) space or network ranges is used to describe the IP address space used by each VPN customer.

The term "Customer Premises Equipment" (CPE) defines an edge device (e.g., router, etc.), fully managed by the provider, connecting a customers PLS to its VPN.

The term "Provider Address" (PA) space or network ranges is used to describe the provider allocated IP addresses in his IP backbone. (e.g., Tunnel endpoints have an address assigned out of the PA range).

The term "Physical LAN Segment" (PLS) is used in this document to describe a broadcast domain, like a shared or switched ethernet segment, connecting hosts, servers and routers at each site. Without the use of a VPN technology, the scope of these PLS is limited per site.

A Virtual Private LAN Segment (VPLS) is the emulation of a LAN segment using Internet facilities. A VPLS can be used to provide what is sometimes known as a transparent LAN service, which can be used to interconnect multiple CPE nodes. It can be seen as a pure layer 2 bridged VPN solution.

The term virtual private networks (VPN) is widely used as a common description for any kind of network built over another network with limited scope.

4

The term "Unnumbered VPN IP" (UVIP) interface is used in VPLS to describe the tunnel endpoint connecting a PLS on a first site with all other PLS per VPN. In the scope of the customer's PLS, this interface doesn't need to have an IP address assigned to forward traffic (VPLS is a layer 2 VPN solution). The tunnel endpoint itself must have an IP

5   address assigned, out of the providers address space.

DETAILED DESCRIPTION OF THE INVENTION

In order to take advantage of all the features of the present invention, it is assumed

10   that the providers of IP backbone services are IP multicast capable. Similarly, it is assumed that CPE devices are able to join a multicast group using IGMP. It is not a requirement that all routers in the backbone have multicast capabilities. It is possible to interconnect the CPE devices via a partially meshed or "star-like" multicast backbone, built using a mix of multicast routing protocols and tunnels to interconnect multicast islands. IP multicast is

15   used to forward broadcast and multicast traffic and for IP address resolution, but not for forwarding of unicast traffic.

Referring now to Fig. 1a, we have shown the physical view or service provider's view of a Virtual Private LAN Segment (VPLS). The IP backbone 10 and CPE devices 11,

20   12, 13 and 14 are managed and typically owned by the service provider. CPE devices 11-14 are typically comprised of routers, whereas each PLS is typically comprised of several IP capable devices such as end stations (ES1, ES2, etc.)

Fig. 1b is a diagram illustrating a logical view of the network of Fig. 1a or as would

25   be seen from the customer's perspective. Whereas in Fig. 1a the CPE devices are visible from the provider's perspective, LAN segments are transparent to the customers as illustrated in Fig. 1b. Similarly, CPE devices which are seen by the service provider are invisible to the customer.

30   Stateless tunnels or links are used in CPE (Customer Premises Equipment) between connected sites. The remote tunnel endpoint address information is directly mapped into the link layer address. ME-ARP is used for IP address resolution inside a VPLS. As a result, VPN connected IP devices will keep all relevant information about the destination tunnel endpoint and VPN membership in their own address resolution (ARP) table.

5

Special unnumbered IP LAN interfaces will generate the link layer address based on a configured VPN identifier and dynamically learned tunnel endpoints (via ME-ARP).

Again, as illustrated in Fig. 1a and 1b, a VPLS can span two or more sites, with all
5   IP devices sharing the same IP subnet. The IP address and mask are chosen by the customer without any restrictions in relation to the provider or other customers. The CPE devices, managed by the provider, are transparent to the customer. This type of layer 2 VPN solution possesses the following benefits for the customer:

10   +   Transparency. No IP addresses must be given to the provider;

            +   Flat IP subnet. The VPN can be seen as a VPLS, with transparent support
                  for broadcast protocols like DHCP/BOOTP (Dynamic Host
                  Configuration Protocol / BOOTstrap Protocol), Netbios/IP etc; and
15

            +   Broadcast and Multicast support. The customer can extend the VPN with
                  their own routers and run any routing protocol over the VPN without any
                  coordination with the provider.

20   Each VPLS has a provider wide unique IP multicast address assigned. A UVIP interface of a CPE device, shown at reference numerals 15, 16, 17 and 18, configured for a particular VPLS, will join the VPN's multicast group by using IGMP. All broadcast traffic is then encapsulated and forwarded to the VPN's IP multicast address. There is therefore no need for a central database to keep track of all UVIP interfaces joining a customer's
25   VPN. This is handled by the IP multicast membership.

            In order to forward IP unicast traffic, an enhanced version of proxy ARP is used. The differences from the standard proxy ARP are:

30   a)   all ARP requests matching the customers IP subnet are encapsulated and
                  forwarded to all VPN members by sending them to the VPN's IP multicast
                  address. Note: The CPE device cannot determine, if an IP device is
                  connected to the local physical segment or not.

35   b)   a forwarded ARP request, after decapsulation, will replace the source
                  hardware address (MAC, Media-Access-Control or physical Address) not

6

with the routers own interface MAC address, but by a calculated address
containing the tunnel source IP address, an interface unique VPN Id (e.g.
VPN instance Id) and a CPE Id (to avoid loops in case of CPE
redundancy).

5

The result of this "multicast enhanced ARP" (ME-ARP) process is that the
customers IP devices will keep all relevant information about the destination tunnel
endpoint and VPN membership in their ARP table. There is no overhead involved, if
compared to a real physical IP subnet.

10

Unique VPN Identifier

Each VPN has a unique identifier assigned. For VPLS built of more than two
physically separated sites this is a valid IP multicast address. As each VPN has a unique IP
15   multicast Id assigned, IGMP and any multicast capable routing protocol (DVMRP
(Distance Vector Multicast Routing Protocol), MOSPF (Multicast Open Shortest Path
First), PIM (Protocol Independent Multicast), are used by a configured IP VPN interface
connecting a Physical Segment to join the VPNs multicast group.

Individual CPE devices are configured as follows:

Based on the VPLS membership using IP multicast, there is no need for a central VPN membership database or protocol to distribute this information. It is enough to
5   configure a new VPN member (physical segment) in the connecting CPE device. The following minimal information is configured per UVIP (Unnumbered VPN IP) interface:

a)   VPN IP multicast Id;

10   b)   IP Network/Mask. Assigned by the customer from the Client Address
                (CA) space. This information is used to determine the correct VPN, based
                on either source or destination IP address. This is important to support
                multi-netting on the same physical interface with many VPNs;

15   c)   Tunnel IP address. This address from the Provider Address (PA) space is
                used to forward VPN traffic over the IP backbone to the correct tunnel
                end-point (bound to a VPN interface). The VPN identifier in each
                encapsulated packet can be used to identify the correct logical UVIP
                interface inside the CPE device;

20

d)   MAC calculation algorithm. This optional, but recommended,
                configuration parameter allows the support of different MAC address
                calculation to prevent possible duplicates.

25   Referring now to Figs. 2a and 2b, in the preferred embodiment of the invention, depending on the security requirements, three different encapsulation formats can be used: without security, with authentication only or with encryption. The encapsulated methods are based on IPsec tunnel mode [RFC2401...RFC2406]. The IP2 header contains the IP source and destination address from the providers address space (tunnel endpoint IP
30   addresses or address as destination address). The IP1 header is the original IP packet header.

In Fig. 2a, we have shown an IPsec AH encapsulation (with authentication). Fig. 2b shows an IPsec ESP encapsulation (with auth. privacy).

35

8

IP multicast and broadcast packets are encapsulated and tagged with the VPN multicast Id in the SPI field of the IPsec AH/ESP header and forwarded to the VPN IP multicast address (equal to VPN multicast Id). All active members of the VPNs multicast group receive the encapsulated packet and forward it to the appropriate VPN's UVIP

5   interface.

Referring now to Figs. 3, we have shown an ARP Request/Reply packet including Ethernet transmission layer. In Fig. 4, we have shown a block diagram of an IP Backbone network and in Fig. 5, we have shown a block diagram illustrating the transfer of packet

10   information between a first and second end station, respectively.

In operation, with reference to Figs. 3, 4, 5 and 6, end station A wants to send an IP packet to end station B on the same logical subnet but connected to different gateways. It is assumed, that the ARP tables 80 and 81 from both end stations are empty. Therefore

15   end station A sends an ARP request 50 to the ethernet broadcast address 51. CPE A, configured with the proper VPN information, checks the source IP address 52 of the ARP request packet 50 against its UVIP interfaces configured on the physical interface. In case of a match, it encapsulates the whole, unmodified, ARP request 50 into an IPsec packet 55 including the VPN identifier 56(equals assigned IP multicast address) and forwards packet

20   55 to the VPN's multicast address 57 using the configured local IP tunnel-endpoint 58 as source address. CPE A also adds a local ARP entry for end station A in its ARP table 72 for that UVIP interface. (CPE A will forward the ARP request, even if end station B is connected to the same physical network).

25   All CPEs joining the VPN will receive this encapsulated ARP request, unpack it, and forward out the local UVIP interface with the following modification to the original ARP request 55:

replace the original HW source address 59 (MAC address from end station A) with

30   a calculated MAC address containing the tunnel end-point IP address from CPE A(= source address from the received IPsec packet) and an optional interface unique VPN Id.

9

This new HW source address 60 is replaced in the ethernet header as well as in the ARP packet 61.

CPE B might add an entry to its ARP table 83 for caching. End station B receives
5   the ARP request 62 and respond to it with a normal ARP reply containing its physical HW MAC address 64 as source in the ethernet header and in the ARP reply packet 65. An ARP entry for end station A with the source MAC address from the ARP request is added on end station B. The ARP table 81 of end station B now contains an entry for end station A with a constructed MAC address containing the tunnel-endpoint IP address and VPN Id.
10   CPE B, configured to listen for constructed MAC addresses, identifies the ARP reply 63 from end station B by checking the source MAC address 64 as well as the source IP address 66 (part of VPN's IP network), encapsulate and forwards the ARP reply 67 directly to the addressed tunnel endpoint (extract tunnel endpoint IP address from destination MAC address).
15

CPE A decapsulates the ARP reply packet 67, checks the destination or target IP address 68 and replaces the destination or target MAC address 69 with the address found in its local ARP cache, and sends the constructed ARP reply 70 out to end station A on the local attached physical LAN segment. In addition, the source MAC address 71(in the
20   Ethernet header and ARP packet) is replaced with a constructed MAC address 72 containing an optional interface locally unique VPN Id and the IP address of CPE B (where the ARP reply came from).

If the ARP table 82 from CPE A does not contain an entry for end station A, then
25   CPE A will have to send an ARP request out for end station A with end station B's IP address before forwarding the ARP reply packet out to end station A.

Finally, end station A receives the ARP reply packet 70 and builds an entry in its ARP table 80 with an entry for end station B and the MAC address containing the remote
30   tunnel endpoint IP address and VPN Id.

10

THE INVENTION CLAIMED IS:

1. A method of sending a unicast IP packet from a first end station to a second end station, said first and second end stations being on the same logical subnet and connected to different CPEs, comprising:

    receiving said unicast IP packet at a CPE associated with said second end station; and

    said CPE associated with said second end station providing said second end station with address resolution information containing mapping information between IP and lower layer physical addresses of said first and second end stations, said lower layer physical addresses being constructed by said CPE and containing VPN membership and physical remote location information such that the constructed lower layer addresses contain enough information for said CPE to forward the packet to the correct remote physical location.

2. A method of sending a multicast IP packet from a first end station to multiple end stations, said first and multiple end stations being on the same logical subnet and connected to different CPEs, comprising:

    receiving said multicast IP packet at each CPE;
    encapsulating said IP multicast packet; and
    forwarding said encapsulated IP multicast packet to a VPN assigned multicast address wherein said IP multicast packet is received by each CPE which has been configured to said VPN.

3. A method as defined in claim 2, wherein said multicast IP packet comprises an IP broadcast packet.

4. A method as defined in claim 2, wherein each of said CPE is configured to said VPN using an IP multicast protocol.

11

5. A method as defined in claim 4, wherein said IP multicast protocol comprises one of an IGMP, DVMRP, MOSPF, MBGP and PIM multicast protocols.

6. A method of sending an IP packet from a first end station to a second end station, wherein said first and second end stations are on the same logical subnet but connected to different CPEs, comprising:

a) sending from said first end station an ARP request with an ethernet broadcast address;

b) at a first CPE associated with said first end station, intercepting said ARP request packet and verifying the intercepted IP address against a corresponding unnumbered virtual packet network (UV) IP interface;

c) if a match is verified, encapsulating said ARP request into an IPsec packet with a VPN identifier; and

d) forwarding said IPsec packet to a VPN's multicast address using configured local IP tunnel-endpoint as a source address.

7. A method as defined in claim 6, wherein said first CPE further adds a local ARP entry for said first end station in its ARP table for said UVIP interface.

8. A method as defined in claim 7, wherein said encapsulated ARP request is received at each CPE connected to said VPN.

9. A method as defined in claim 8, wherein said ARP request is unpacked, modified and forwarded out of the local UVIP interface when received at said CPE.

10. A method as defined in claim 9, wherein said ARP request is modified at each CPE by replacing the original HW source address with a calculated MAC address containing the tunnel end-point IP address from said first CPE and an interface unique VPN Id thus providing a new HW source address to replace in the ethernet header as well as in the ARP packet itself.
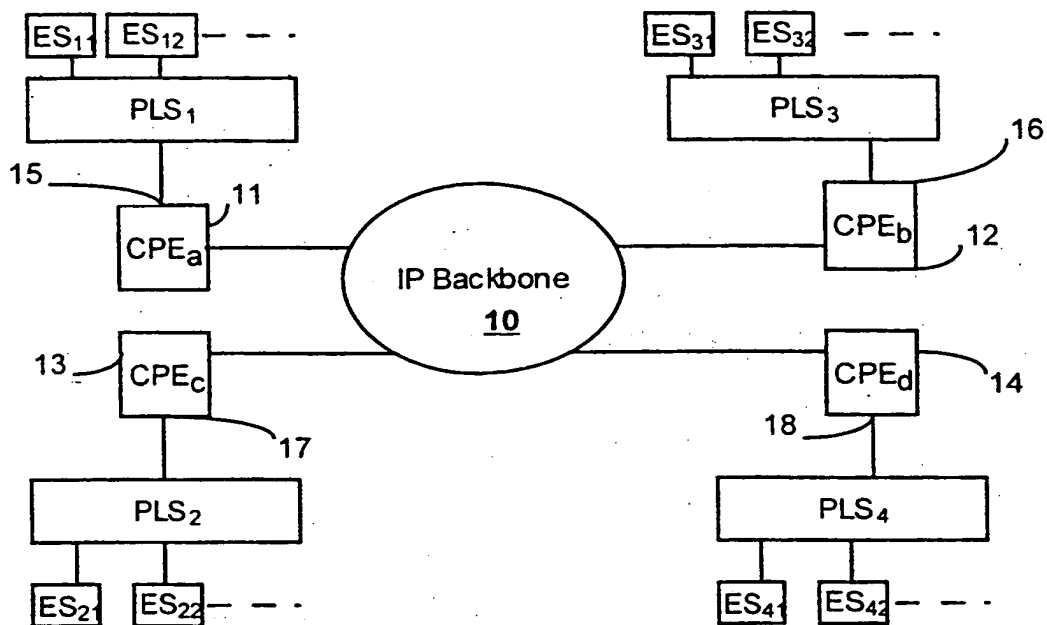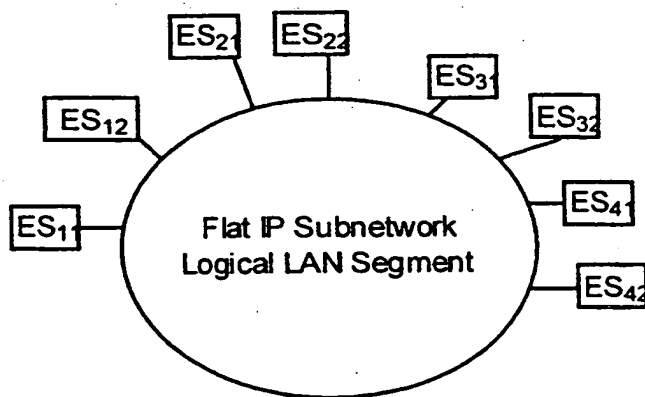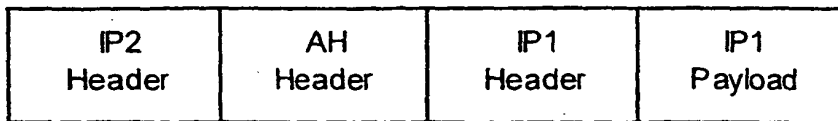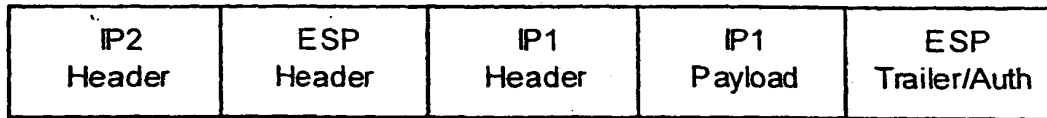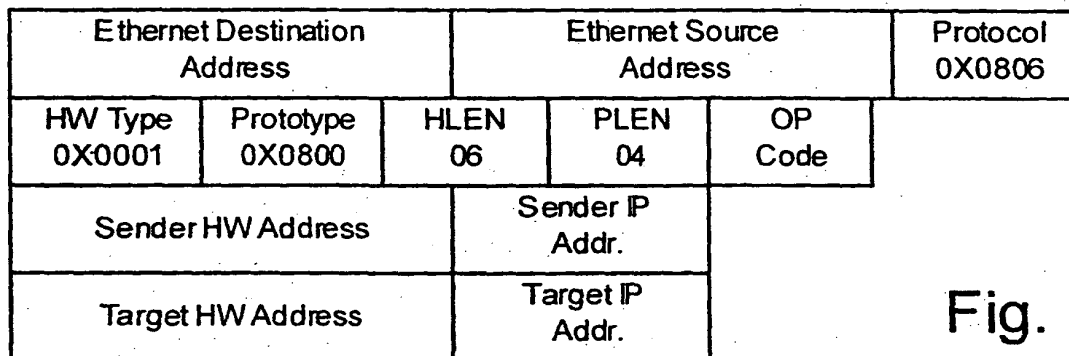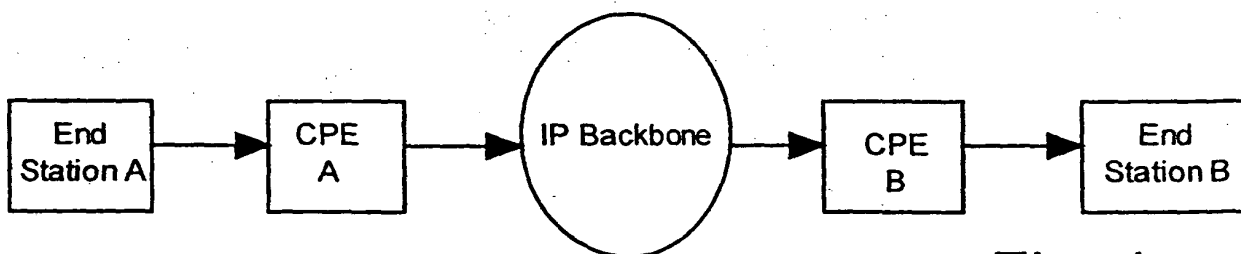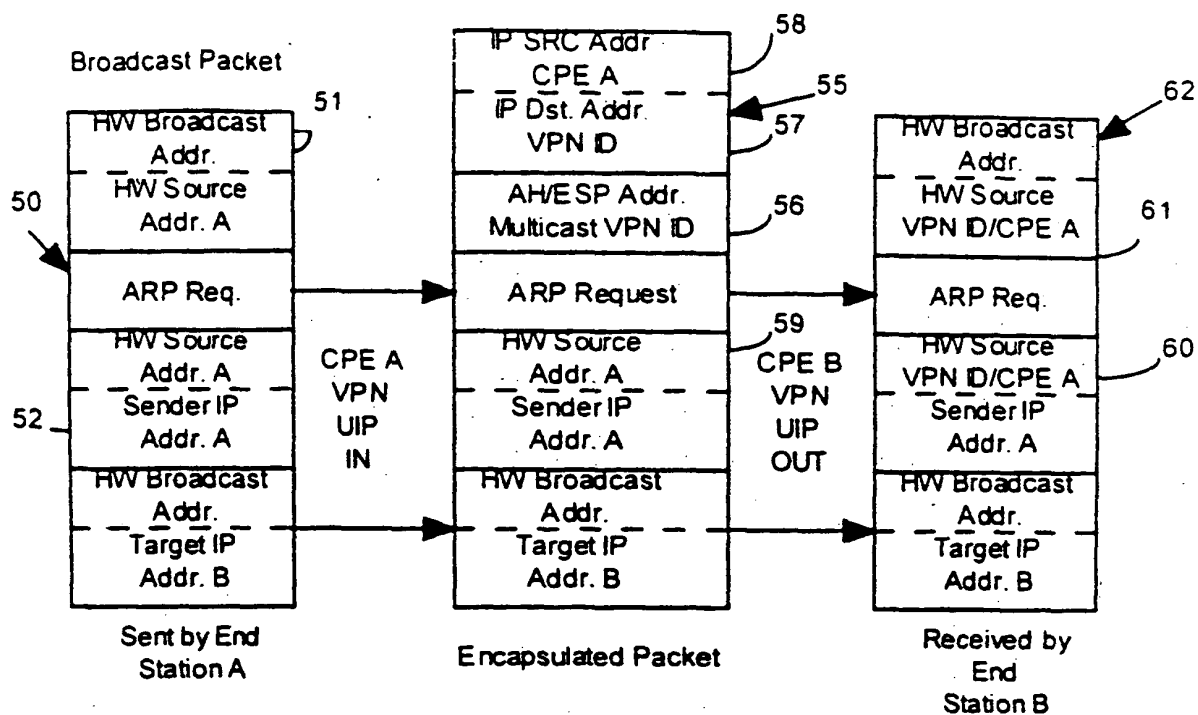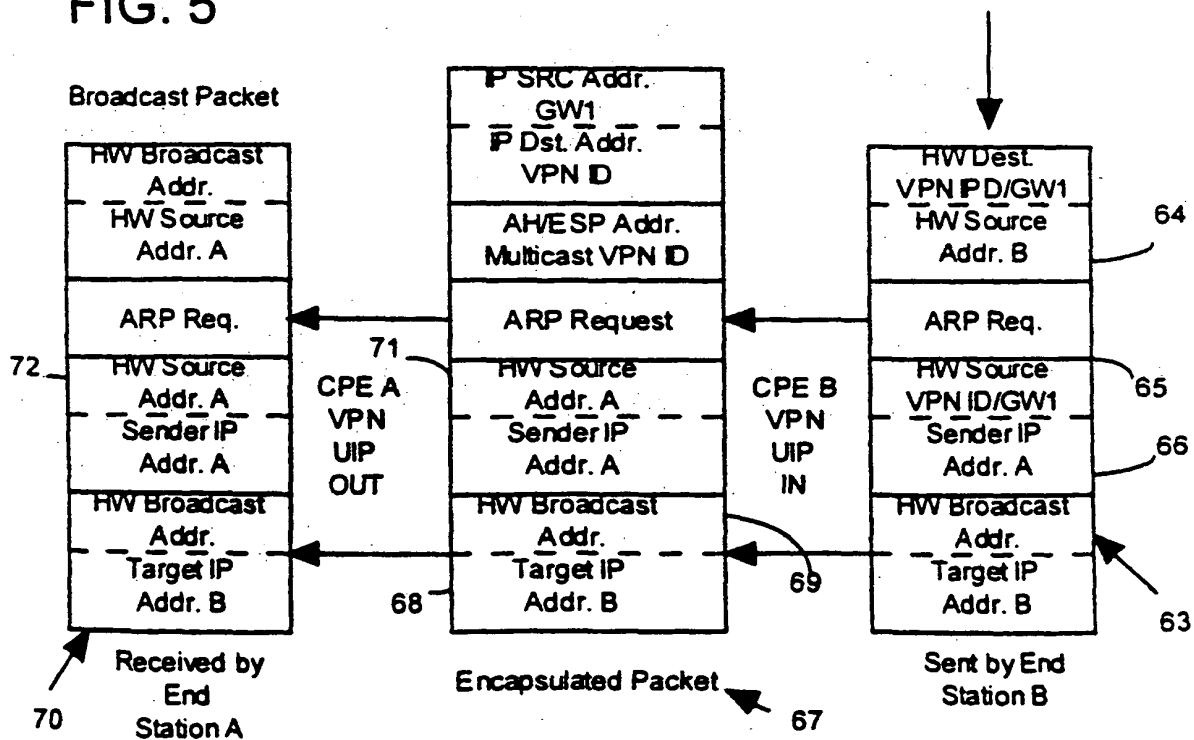
12

Fig. 1a

Fig. 1b

IP2 Header	AH Header	IP1 Header	IP1 Payload

Fig. 2a

IP2 Header	ESP Header	IP1 Header	IP1 Payload	ESP Trailer/Auth

Fig. 2b

Ethernet Destination Address		Ethernet Source Address		Protocol 0X0806
HW Type 0X0001	Prototype 0X0800	HLEN 06	PLEN 04	OP Code
Sender HW Address		Sender IP Addr.		
Target HW Address		Target IP Addr.		

Fig. 3

End Station A	→	CPE A	→	IP Backbone	→	CPE B	→	End Station B

Fig. 4

3/3

FIG. 5

FIG. 6

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7   H04L12/46   H04L12/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 ` H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, INSPEC, COMPENDEX, IBM-TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 98 02821 A (3COM CORP) 22 January 1998 (1998-01-22) abstract page 4, line 10 -page 6, line 29 page 13, line 9 -page 15, line 29	1-10
A	EP 0 812 086 A (NIPPON TELEGRAPH & TELEPHONE) 10 December 1997 (1997-12-10) page 2, column 2, line 37 -page 3, column 3, line 35 page 4, column 5, line 2 -page 11, column 50	1-10
A	WO 98 57465 A (VPNET TECHNOLOGIES INC) 17 December 1998 (1998-12-17) page 3, line 13 -page 4, line 8 page 9, line 1 -page 11, line 8	1-10

-/--

[X] Further documents are listed in the continuation of box C.   [X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 August 2000	04/09/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Karavassilis, N

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	"VIRTUAL PRIVATE NETWORKS ON VENDOR INDEPENDENT NETWORKS" IBM TECHNICAL DISCLOSURE BULLETIN,US,IBM CORP. NEW YORK, vol. 35, no. 4A, 1 September 1992 (1992-09-01), pages 326-329, XP000314784 ISSN: 0018-8689 page 327, line 12 -page 329, column 24	1-10

1

INTERNATIONAL SEARCH REPORT

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9802821	A	22-01-1998	US	6041166 A	21-03-2000
			AU	3728497 A	09-02-1998
			GB	2330285 A	14-04-1999
ER 0812086	A	10-12-1997	JP	10056473 A	24-02-1998
			US	6075776 A	13-06-2000
WO 9857465	A	17-12-1998	AU	7837998 A	30-12-1998
			EP	09887 A	29-03-2000